



Cosmological Constraints on Pseudo-Nambu-Goldstone Bosons

Joshua A. Frieman

*NASA/Fermilab Astrophysics Center
Fermi National Accelerator Laboratory
Batavia, IL 60510-0500*

and

Andrew H. Jaffe

*Department of Astronomy and Astrophysics
The University of Chicago
5640 South Ellis Avenue, Chicago, IL 60637*

Abstract

Particle physics models with pseudo-Nambu-Goldstone bosons (PNGBs) are characterized by two mass scales: a global spontaneous symmetry breaking scale f and a soft (explicit) symmetry breaking scale Λ . We investigate general model-insensitive constraints on this two-dimensional parameter space arising from the cosmological and astrophysical effects of PNGBs. In particular, we study constraints arising from vacuum misalignment and thermal production of PNGBs, topological defects, and the cosmological effects of PNGB decay products, as well as astrophysical constraints from stellar PNGB emission. Bounds on the Peccei-Quinn axion scale, $10^{10} \text{ GeV} \lesssim f_{PQ} \lesssim 10^{10} - 10^{12} \text{ GeV}$, emerge as a special case, where the soft breaking scale is fixed at $\Lambda_{\text{QCD}} \simeq 100 \text{ MeV}$.



I. Introduction

Particle physics models with spontaneously broken global symmetries are quite common; by Goldstone's theorem, the spectrum of such theories must contain a massless spin-0 boson for each broken symmetry group generator. If the global symmetry is only approximate, *i.e.*, it is explicitly (in addition to being spontaneously) broken, the associated bosons become massive pseudo-Nambu-Goldstone bosons (PNGBs). In nature, the best known example of this phenomenon is the π meson, associated with chiral symmetry breaking. Particle theory provides a host of additional PNGB candidates,¹ including axions, majorons, familons, and schizons (although the majoron and familon may be exactly massless Nambu-Goldstone bosons).

Such models are generally characterized by two mass scales: the spontaneous symmetry breaking scale f , and an explicit breaking scale Λ . The simplest example is that of a complex scalar Φ with a potential $V(\Phi^*\Phi)$ which is minimized at $\Phi = fe^{i\phi/f}$. The non-zero vacuum expectation value of Φ spontaneously breaks the global $U(1)$ symmetry $\Phi \rightarrow \Phi e^{i\alpha}$ at the scale f , and the angular field ϕ is the massless Nambu-Goldstone mode around the bottom of the 'Mexican hat' potential. At the lower scale Λ , a periodic potential for ϕ of height $\sim \Lambda^4$ is generated. (The form of the potential may reflect a residual discrete symmetry.) The resulting PNGB has a mass given by $m_\phi \sim \Lambda^2/f$. In models with a large hierarchy between the scales f and Λ ($f \gg \Lambda$), PNGBs are thus very light and also very weakly interacting, since their couplings are suppressed by inverse powers of f . Nevertheless, they can play an important role in astrophysics and cosmology. For example, in axion models, where f is the Peccei-Quinn scale f_{PQ} and the symmetry is explicitly broken by QCD instantons (through the chiral anomaly) at the chiral symmetry breaking scale $\Lambda_{\text{QCD}} \simeq 100$ MeV, cosmological² and astrophysical³ arguments constrain f_{PQ} to lie in a narrow window around $f_{PQ} \simeq 10^{10}$ GeV (perhaps extending up to 10^{12} GeV, although this point is controversial; see below).

Although motivated by the strong CP problem, the QCD axion is a particular instance of a more general phenomenon. For example, Hill and Ross⁴ have explored schizon models, in which Λ is associated with the mass of a light fermion (quark or lepton), and might plausibly lie in the eV (for neutrinos) to MeV (for charged leptons or quarks) range. On the other hand, superstring models contain one or more very heavy axion fields; for example, for the model-independent axion⁵ (the imaginary counterpart to the dilaton), Λ is associated with the scale at which the gauge coupling in the hidden sector group becomes large. This naturally happens at a very high energy scale,⁶ $\Lambda \sim 10^{14} - 10^{17}$ GeV. (Indeed, it has been suggested that a PNGB with $f \sim m_{\text{pl}}$ and $\Lambda \sim 10^{15}$ GeV is a natural candidate for the inflaton field.⁷)

The lesson we draw from this is that there are a number of models with axions or axion-like particles in which the scale Λ has no connection with the QCD scale; indeed, from above we see that it may vary from the atomic to the Planck scale. We are thus led to consider the more general phenomenon of PNGB models in which both f and Λ are *a priori* unconstrained. In this paper, we investigate the cosmological and astrophysical constraints on this two-dimensional parameter space. These bounds have previously been studied in detail for the QCD axion with fixed $\Lambda = \Lambda_{\text{QCD}}$; here, we explore how these constraints are altered when the scale Λ is allowed to vary over a wide range. In Sec. II, we discuss PNGB production from initial vacuum angle misalignment and the bound which results from requiring that the PNGB density satisfy $\Omega_\phi h^2 \leq 1$. In Sec. III, we consider PNGB production from topological defects (strings, textures, and global monopoles) and, more generally, from long wavelength spatial gradients arising from the finite correlation length of the scalar field. We consider thermal production of PNGBs in Sec. IV. In Sec. V, we study a new wrinkle in PNGB phenomenology not present in ‘invisible’ axion models: if they are sufficiently massive, PNGBs decay on a timescale shorter than the age of the Universe; we examine the concomitant constraints on their decay products. In Sec. VI, we discuss astrophysical bounds arising from PNGB emission from red giants and supernova 1987a, and we conclude in Sec. VII. In all cases, the bounds we obtain are relatively model-insensitive: aside from order unity couplings, the form of the PNGB interactions are fixed by symmetry, so they are not strongly model-dependent.

Here, we briefly define some notation. In general, the PNGB ϕ will be taken to have a (low-temperature) mass given by

$$m_\phi = \frac{\Lambda^2}{f} = 10^{-5} \text{ eV} \left(\frac{\Lambda}{100 \text{ MeV}} \right)^2 \left(\frac{10^{12} \text{ GeV}}{f} \right). \quad (1.1)$$

Eqn.(1.1) may be thought of as defining the scale Λ in terms of the PNGB mass and f . In general, the PNGB field is related to an angular parameter of the broken symmetry, $\phi = f\theta$, where θ runs from 0 to 2π : we have absorbed any “winding” of the field into f .

II. Non-Thermal Misalignment Production

At high temperatures ($T \gg \Lambda$), the angular degree of freedom representing the PNGB is randomly oriented. As the universe cools below a critical temperature T_1 , the ϕ field will roll to the minimum of its potential and will begin to oscillate coherently, resulting in a Bose condensate that can be treated as a classical field configuration.

Expanded about the minimum of the potential, the Lagrangian density of the PNGB can be written as

$$\mathcal{L} = \frac{1}{2} \dot{\phi}^2 - \frac{1}{2} m_\phi(T)^2 \phi^2 \quad (2.1)$$

for a spatially homogeneous field ϕ , to quadratic order in the field. (We consider the effects of inhomogeneities in the field in the next section.) In general, the potential $V(\phi)$ will have higher order terms, but these are comparatively small for $\phi \lesssim f$ ($\theta \lesssim 1$); these anharmonic effects in the potential will slightly increase the PNGB energy density⁸ if $\theta \gtrsim 1$.

We will parametrize the temperature-dependent mass of a general PNGB as

$$m_\phi(T) = \alpha m_\phi(T=0) \left(\frac{\Lambda}{T}\right)^\nu \quad T \gtrsim \Lambda \quad (2.2)$$

where α and ν are fitting parameters of order unity. In the case of the QCD axion, the PNGB gains a mass at finite temperature due to QCD instantons⁹; in this case, the behavior is approximately⁸ $\alpha \simeq 0.1$ and $\nu \simeq 3.7$. However, the form of Eq. (2.2) holds more generally; the temperature-dependent mass in schizon models has also been discussed recently¹⁰.

The classical equation of motion for ϕ is as usual for a scalar field in an expanding spacetime (ignoring for now the decay width of ϕ):

$$\ddot{\phi} + 3H\dot{\phi} + m_\phi(T)^2\phi = 0. \quad (2.3)$$

At high temperatures where $m_\phi(T) \ll 3H(T)$, $\phi = \phi_1 = f\theta_1$ is a constant and frozen to its initial value. As the temperature decreases, the mass turns on adiabatically; the field begins to oscillate coherently at a temperature T_1 given by $m_\phi(T_1) = 3H(T_1)$. The PNGB energy density, averaged over an oscillation, then scales as

$$\rho_\phi = \langle \dot{\phi}^2 \rangle \propto m_\phi(T)/R^3, \quad (2.4)$$

where $R(t)$ is the scale factor of the universe; thus the number of PNGBs per comoving volume is conserved, $n_\phi = \rho_\phi/m_\phi(T) \propto R^{-3}$. [At low temperatures, when $m_\phi(T) = m_\phi(0) = \text{const.}$, the PNGB energy density scales like non-relativistic matter, $\rho_\phi \propto R^{-3}$ ($T \ll \Lambda$).] As long as the expansion of the universe is adiabatic, the entropy density s is also proportional to R^{-3} , so the ratio n_ϕ/s has been constant since the onset of coherent PNGB oscillations. If the entropy in a comoving volume has increased by a factor γ since the onset of PNGB oscillations, the ratio n_ϕ/s will simply be reduced by this factor. Thus we can find the present PNGB abundance by calculating the PNGB-to-entropy ratio when coherent oscillations begin; since $\rho(T_1) = m_\phi^2(T_1)\phi_1^2/2$, we have

$$\left. \frac{n_\phi}{s} \right|_{T_1} \simeq \frac{m_\phi(T_1)\phi_1^2/2}{2\pi^2 g_{*1} T_1^3/45} \quad (2.5)$$

where $s = 2\pi^2 g_* T^3/45$ is the entropy density, $g_*(T)$ counts the number of relativistic degrees of freedom at temperature T , and the subscript 1 denotes the value at T_1 .

Reexpressing this in terms of the initial misalignment angle θ_1 and the low-temperature mass m_ϕ , using $m_\phi(T_1) = 3H_1 = 5g_{*1}^{1/2}T_1^2/m_{\text{pl}}$, appropriate for the radiation-dominated early universe, we can calculate the present PNGB density in terms of the critical density (defined by $H_0^2 = (8\pi G/3)\rho_{\text{crit}}$)

$$\begin{aligned}\Omega_{\phi,\text{mis}} &\equiv \frac{\rho_\phi}{\rho_{\text{crit}}} = \frac{(n_\phi/s)(s_0/\gamma)m_\phi}{\rho_{\text{crit}}} \\ &= 1.3 \times 10^{-10} h^{-2} \text{GeV}^{-2} \frac{\Lambda^4 \theta_1^2 \gamma^{-1}}{g_{*1} m_\phi T_1}.\end{aligned}\quad (2.6)$$

Here, the present Hubble parameter is $H_0 = 100h \text{ km/sec/Mpc}$, and observations indicate $0.4 \lesssim h \lesssim 1$; we have also used the fact that the present entropy density is $s_0 = 2970 \text{ cm}^{-3}$ (for a photon temperature¹¹ of 2.735 K). Finally, T_1 is given by $m_\phi(T_1) = 3H(T_1)$; for reference, in a radiation-dominated universe,

$$\frac{T_1}{\Lambda} = \left(\frac{5g_{*1}^{1/2} f}{\alpha m_{\text{pl}}} \right)^{\frac{-1}{\nu+2}}. \quad (2.7)$$

Substituting Eq. (2.7) into (2.6), we find

$$\Omega_{\phi,\text{mis}} h^2 = 1.3 \times 10^{-10} \theta_1^2 \left(\frac{\Lambda f}{\text{GeV}^2} \right) \left(\frac{f}{0.2\alpha m_{\text{pl}}} \right)^{\frac{1}{\nu+2}} \gamma^{-1} g_{*1}^{-\frac{2\nu+3}{2\nu+4}} \quad (2.8)$$

This expression merits several comments. First, it retains some implicit dependence on T_1 through the term in g_{*1} ; however, between $T \sim 1 \text{ TeV}$ and today, g_* has only changed by about two orders of magnitude, so this additional parameter does not have a substantial effect. Second, we have assumed in Eqns. (2.6) and (2.7) that the universe is radiation-dominated at T_1 . Thus, Eqn. (2.8) is only strictly valid if $m_\phi > 3H(T_{\text{eq}}) \simeq 10^{-26} \text{ eV}$, where T_{eq} is the temperature when the universe first becomes matter-dominated. Third, we note that for the axion, the most well-studied and well-understood PNGB, there is a theoretical uncertainty in this value of about⁸ $10^{\pm 0.4}$, coming mostly from the parametrization of the finite temperature mass; for the general case we might expect this to be larger.

Solving this for Λ , and assuming from here on negligible entropy production, we find

$$\frac{\Lambda}{\text{GeV}} = \left[(2.3 \times 10^9)(2.4 \times 10^{18})^{\frac{1}{\nu+2}} \right] g_{*1}^{\frac{2\nu+3}{2\nu+4}} \alpha^{\frac{1}{\nu+2}} \left(\frac{\theta_1}{\pi/\sqrt{3}} \right)^{-2} \left(\frac{f}{\text{GeV}} \right)^{-\frac{\nu+3}{\nu+2}} \Omega_{\phi,\text{mis}} h^2. \quad (2.9)$$

We can use our knowledge of the expansion and age of the universe to restrict these parameters: since the Hubble parameter $h > 0.4$ and the age of the universe $t_0 > 10^{10} \text{ yr}$ (from globular cluster and nucleocosmochronology dating) we have the constraint on the cosmic density,¹² $\Omega_0 h^2 \leq 1$. This requirement translates into an allowed region of the Λ - f

plane, for fixed values of the $O(1)$ parameters (*i.e.*, $\alpha, \nu, g_{*1}, \gamma, \theta_1$). We show this region in Fig. 1.

We have chosen as a typical value of the initial misalignment angle $\theta_1 = \pi/\sqrt{3}$ —the RMS value of an angle randomly distributed between $-\pi$ and π . On average, we expect this value to hold because the present Hubble volume comprises a very large number of regions that were causally disconnected at $t_1 = t(T_1)$, within each of which θ_1 was arbitrary. However, if the universe inflated during or after the PNGB symmetry breaking (at $T \sim f$) without reheating to temperatures above the symmetry-breaking scale f , we have no way of knowing the initial angle in our observable neighbourhood. In an inflationary scenario, it may be possible to escape this bound entirely if $\theta_1 \simeq 0$ in our neighbourhood of the universe, since the observable Hubble volume would come from a single causally connected region within which θ_1 was homogeneous. See refs. 13 for discussion of these issues.

Figure 1 shows that this process alone excludes a large portion of the Λ – f plane above the line representing $\Omega h^2 = 1$. Note the spread of 1–2 orders of magnitude by varying the model-dependent parameters about the “axionic” case. In particular, for small values of θ_1 , the amount of excluded parameter space is decreased rapidly since $\Omega_{\phi, \text{mis}} \propto \theta_1^2$.

We have also shown, for reference, lines of constant mass m_ϕ . For the axion, located on the horizontal line at $\Lambda = \Lambda_{\text{QCD}} \simeq 100 \text{ MeV}$, misalignment production requires $f_{\text{PQ}} \lesssim 10^{13} \text{ GeV}$, in agreement with previous analyses^{2,8}.

III. PNGB Production from Symmetry Breaking

When the universe cools below $T \sim f$ and the symmetry is spontaneously broken, global defects appropriate to the topology of the vacuum manifold will form¹⁴. In the simplest case (which includes axions), the relevant broken symmetry is $U(1)$, the vacuum manifold S^1 has non-trivial first homotopy group, $\pi_1(S^1) = \mathbb{Z}$ (the integers), and the defects produced are global strings; we first consider this case.

At $T \sim f$, the massive complex field $\Phi = \sigma e^{i\phi/f}$ rolls down to its minimum, $|\Phi| = \sigma = f$, anywhere around the “mexican hat,” since the potential for the phase of Φ (the PNGB field ϕ) does not have discrete minima until $T \sim \Lambda$. Because the angular field ϕ is initially uncorrelated on scales larger than the horizon, there will be closed loops in space around which ϕ winds by $2\pi N$, where N is an integer; by continuity, there must be some point within each surface spanned by such a loop where the angular field is undefined. These points are the cores of global strings, where the complex field is trapped in the false vacuum, $\Phi = 0$. Due to the gradient and potential energy of the configuration, these strings have an energy per unit length $\mu \simeq \pi f^2 \ln(fd)$, where d is a characteristic distance between strings. (The logarithm comes from the gradient energy in the field.)

In the absence of inflation, the global string network should, as in the case of gauge

strings, quickly reach the so-called scaling solution¹²; for long strings, the energy density then is of order

$$\rho_{\text{string}} \sim \mu t^{-2} \quad \frac{\rho_{\text{string}}}{\rho_{\text{total}}} \sim G\mu. \quad (3.1)$$

For the scaling solution to be maintained, almost all the energy in strings must be radiated away each Hubble time as PNGBs (this is the crucial difference between global strings and gauge strings, which primarily decay through gravitational radiation). Thus the change in the relative PNGB abundance over one Hubble time will be

$$\Delta(n_\phi/s) \sim \frac{\mu t^{-2}}{\omega T^3} \quad (3.2)$$

where ω is the average energy per radiated PNGB. This process will be effective from the epoch of spontaneous symmetry breaking ($T \sim f$) until the time when the PNGB mass m_ϕ becomes comparable to the expansion rate H (at temperature T_1 as before): the latter epoch corresponds to the time when: (1) the strings become connected by domain walls (at $T \sim \Lambda \sim T_1$) and rapidly chop themselves up, and (2) the strings can no longer radiate into very low-frequency PNGB modes (since $m_\phi \sim H(T_1) \sim \omega(t_1)$) which, at least in the Davis-Shellard scenario (see below), dominate PNGB production. Using the value of the Hubble parameter appropriate to a radiation dominated universe ($H \sim t^{-1} \sim T^2/m_{\text{pl}}$) and integrating Eqn.(3.2) gives

$$\frac{n_\phi}{s} \sim \mu \int_{T_1}^f \frac{dT}{\omega(t)m_{\text{pl}}^2}. \quad (3.3)$$

For global strings, there is considerable controversy surrounding the value of $\omega(t)$, the average energy per radiated PNGB. Davis and Shellard¹⁵ (DS) claim that the PNGBs should be radiated predominantly into low-frequency modes—wavelengths of order the horizon size ($\omega \sim t^{-1}$)—but Harari and Sikivie¹⁶ (HS) argue for a $1/k$ spectrum, giving $\omega \sim \ln(ft)/t$. Rather than make a decision regarding these contentious issues, we will parametrize our ignorance: the density of particles produced by the decay of global strings is given by

$$\frac{n_\phi}{s} \sim S \frac{f^2}{T_1 m_{\text{pl}}} \left(1 - \frac{T_1}{f}\right) \quad (3.4)$$

where the value of S is a result of this debate over the PNGB spectrum. For the DS scenario, $S \sim \ln(ft_1)$ (note that the logarithm comes from $\mu \sim f^2 \ln(fd)|_{t_1} \sim f^2 \ln(ft_1)$); for the HS scenario, $S \sim 1$, since the $\ln(ft)$ in ω (taken out of the integral since it is much more slowly varying than the rest of the integrand) cancels with that from μ . Since

$T_1 \sim \Lambda \ll f$, the energy density of string-produced PNGBs is

$$\begin{aligned}\Omega_{\phi,s} h^2 &= 2.3 \times 10^{-11} \text{ GeV}^{-2} S \frac{\Lambda^4}{m_\phi T_1} \\ &= 2.3 \times 10^{-11} S \left(\frac{\Lambda f}{\text{GeV}^2} \right) \left(\frac{f}{0.2 \alpha m_{\text{pl}}} \right)^{\frac{1}{\nu+2}} g_{*1}^{\frac{1}{2\nu+4}}\end{aligned}\quad (3.5)$$

Aside from the factor of S , the abundance of string-produced PNGBs is comparable to that produced by the misalignment mechanism (compare Eqn.(2.8)), and may complement this mechanism or supplant it. Thus the string-produced density—and the resulting bound on f and Λ —for the HS case is comparable to that for misalignment production.

For the DS scenario, we need to compute the quantity $\ln(ft_1)$; using Eqn.(2.7) we find $ft_1 = (3f^2/2\alpha\Lambda^2)(5g_{*1}^{1/2}f/\alpha m_{\text{pl}})^{-\nu/(\nu+2)}$, or

$$\ln(ft_1) = 93 + 2 \ln \left(\frac{100 \text{ MeV}}{\Lambda} \right) + \left(\frac{\nu+4}{\nu+2} \right) \ln \left(\frac{f}{m_{\text{pl}}} \right) - \left(\frac{\nu}{\nu+2} \right) \ln \left(5g_{*1}^{1/2} \alpha^{2/\nu} \right) \quad (3.6)$$

In solving the equation for Λ , we can neglect the $\ln \Lambda$ term provided $m_\phi \ll m_{\text{pl}}$, valid over most regimes of physical interest. Furthermore, for the axion with $\alpha \simeq 0.1$ and $\nu \simeq 3.7$, the third term has a maximum value of about 2. Thus, we can approximate $\ln(ft_1) \simeq 91 + [(\nu+4)/(\nu+2)] \ln(f/m_{\text{pl}})$ and we can solve for limits on Λ due to cosmic string decay:

$$\frac{\Lambda}{\text{GeV}} = \left[(5.1 \times 10^{10})(2.4 \times 10^{18})^{\frac{1}{\nu+2}} \right] \frac{(f/\text{GeV})^{-\frac{\nu+3}{\nu+2}} g_{*1}^{\frac{-1}{2\nu+4}} \alpha^{\frac{1}{\nu+2}} \Omega_{\phi,s} h^2}{91 + \left(\frac{\nu+4}{\nu+2} \right) [\ln(f/10^{12} \text{ GeV}) - 16]} \quad (3.7)$$

for DS strings; remove the logarithmic factor in the denominator for the HS case. Despite the presence of the logarithm, the dependence of Λ upon f is still approximately that of a power law. If the DS analysis is correct, the PNGB abundance is increased by almost two orders of magnitude and the limits on the parameter space are correspondingly tighter.

If the global symmetry group is not $U(1)$, other defects will form, but similar results will obtain as long as these defects also follow scaling solutions like Eqn.(3.1). For example, if the symmetry breaking is $SU(2) \rightarrow U(1)$, global monopoles will form. Each monopole has an energy $\simeq 4\pi f^2 R$, where R is a characteristic distance between monopoles. Numerical simulations¹⁷ show that there should be $O(1-10)$ global monopoles per Hubble volume, $R \sim t \sim H^{-1}$, so the energy density in monopoles, $\rho \sim (f/t)^2$, obeys the scaling solution. Global textures¹⁸ which form, for example, in the symmetry breaking $O(4) \rightarrow O(3)$ also obey such a solution, and the results above will obtain. In general, then, the defects arising in global spontaneous symmetry breaking give rise to a PNGB density comparable to that produced by the misalignment mechanism, up to factors of order $\ln(ft_1)$ (which depend on details of the spectrum and field configuration for the defect).

The bulk of the energy density of global strings, monopoles, or textures is carried in the gradient energy of the PNGB field itself (unlike gauge strings or monopoles, in which the energy is carried in the core of the defect). Thus, even ignoring the defects themselves, “Kibble” gradients in the PNGB field (which arise from the field taking on different random values in initially causally disconnected regions) will also result in a population of PNGBs: as the horizon grows to encompass previously uncorrelated regions, the PNGB field will smooth itself, with a correlation length $\xi \sim 1/H \sim t$. At temperatures $T \gtrsim T_1$, while the PNGB is still effectively massless, the gradient energy density is

$$\begin{aligned}\rho_{\text{grad}} &\sim \frac{1}{2}(\nabla\phi)^2 = \frac{1}{2}f^2(\nabla\theta)^2 \\ &\sim \frac{1}{2} \frac{f^2 \langle \theta^2 \rangle}{\xi^2} \sim f^2 t^{-2}.\end{aligned}\tag{3.8}$$

Once these long-wavelength fluctuations in ϕ enter the horizon, we can consider them as a coherent state of ϕ particles. Thus, spatial gradients in the PNGB field will produce a PNGB energy density again comparable to that produced by the misalignment mechanism.

The excluded regions of the Λ - f plane are shown in Fig. 2. In this figure and below we only show the line for one set of model-dependent parameters, those appropriate for the axion. Bear in mind that these lines will have a spread of 1–2 orders of magnitude (compare Fig. 1). We show results for both DS and HS strings; we expect that any PNGBs produced from either topological defects or spatial gradients should span this range. For the axion itself, at $\Lambda = \Lambda_{\text{QCD}} \simeq 100 \text{ MeV}$, HS strings restrict $f \lesssim 10^{12} \text{ GeV}$ (not much different than the misalignment production bound), and DS strings restrict $f \lesssim 4 \times 10^{10} \text{ GeV}$ (an improvement of almost two orders of magnitude).

When the soft symmetry breaking that gives the PNGB its mass occurs at $T \sim \Lambda$, further topological defects may form, and we must insure that they are cosmologically benign. In the $U(1)$ case, initially the PNGB field varies smoothly from 0 to 2π around the string. When the symmetry is explicitly broken at $T \sim \Lambda$, the PNGB field is forced to its minimum, but the variation over 2π cannot be removed. Thus, domain walls will form, bounded by strings. If the PNGB potential has a unique (non-degenerate) minimum, each string is bounded by a single wall, and the string-wall system rapidly chops itself up and disappears. (This assumes the strings have not been inflated away.) If the PNGB potential has multiple minima, however, or if the strings are inflated away before the walls form, the domain walls would come to dominate the energy density of the universe and the microwave background would be strongly anisotropic¹⁹. Related phenomena may occur in other scenarios as well. For example, if the initial defects are global monopoles, the second symmetry breaking will result in monopoles connected by strings, leading to rapid monopole annihilation; in this case, the secondary defects are harmless.

IV. Thermal PNGB Production

So far, we have assumed that PNGBs were not in thermal equilibrium in the early universe. That is, we have been assuming that microphysical scattering processes have had small impact on the abundance of PNGBs. Now, we consider the case in which PNGBs may couple to other constituents of the early universe and the conditions under which thermal production may be important.

Since we wish to study constraints on PNGB properties that are as model-independent as possible, we do not know its couplings *a priori*. In general, however, the coupling strength will be suppressed by powers of $(1/f)$. We will consider two possible cases: (I) the PNGB is coupled to ordinary matter (*e.g.*, quarks, leptons, and photons). (II) the PNGB is coupled only to the matter sector associated with the scale Λ , that is, to particles with masses $\tilde{m} = (\Lambda/\Lambda_{\text{QCD}})mf(G)$, where m is the mass of a quark or hadron, and $f(G)$ depends on the gauge group associated with the scale Λ . For example, a “techni-axion” coupled only to technicolor particles would fall in category (II), with Λ the scale at which the technicolor gauge group becomes strong (\sim the electroweak scale).

The interactions of the PNGB with photons and fermions are¹

$$\mathcal{L} = g_{\phi\gamma\gamma}\phi\mathbf{E}\cdot\mathbf{B} + \sum_{\text{fermions}} i\frac{g_{\phi ff}}{2m_f}\partial_\mu\phi(\bar{f}\gamma^\mu\gamma_5 f) \quad (4.1)$$

with coupling constants

$$\begin{aligned} g_{\phi\gamma\gamma} &= g_\gamma \frac{\alpha/2\pi}{f} \\ g_{\phi ff} &= g_f \frac{m_f}{f} \end{aligned} \quad (4.2)$$

Here, g_f and g_γ are constants assumed to be of order one, although in some cases they may vanish (*e.g.*, for the hadronic axion, which does not couple to electrons at tree level, $g_e = 0$) or may be large (*e.g.*, in Sikivie’s omion model²⁰, $g_\gamma \sim 10^5$). In general, the values of these constants depend on the appropriate gauge-group charges, and to first order do not depend on f or Λ . Note that we have assumed the coupling to fermions is purely pseudoscalar, as is the case for axions, majorons, and familons in models without schizons.

We will assume (following ref. 21) that the PNGB is produced in reactions like $ab \leftrightarrow \phi X$ in the early universe, with species a , b , and X all in thermal equilibrium. The PNGB will be produced in copious amounts by scattering processes if it was ever in thermal equilibrium *i.e.*, if its rate of production, Γ , exceeds the expansion rate of the universe, H . In that case, it will have an abundance given by the equilibrium value for relativistic bosons, $n_\phi^{\text{EQ}} \cong 1.2T^3/\pi^2$ corresponding to a present energy density

$$\Omega_{\phi,\text{TE}}h^2 = \frac{m_\phi}{130\text{ eV}} \frac{10}{g_{*F}} \quad (4.3)$$

where g_{*F} is the value of g_* at the PNGB freeze-out temperature T_F (defined by $\Gamma(T_F) = H(T_F)$). Here we have assumed the PNGBs are relativistic at freeze-out, $T_F \gg m_\phi$; since we will see below that $T_F \sim \Lambda$, this just corresponds to the usual assumption that $f \gg \Lambda$.

Thus the crucial model-dependence enters into the calculation of Γ/H to decide if the PNGB has been in thermal equilibrium. If the PNGB couples to normal matter (case (I) above), the calculation is the same as that for the axion, and only the value of f —not Λ —enters. The PNGB abundance is controlled by two processes. Before the quark-hadron transition at $T \sim \Lambda_{\text{QCD}} \sim m_N/5$ (where $m_N \simeq 1 \text{ GeV}$ is the nucleon mass), the dominant production mechanism is PNGB photo- or gluon production in the presence of a heavy, but still relativistic, quark, Q (since the coupling strength is proportional to the fermion mass) $Q\gamma \rightarrow Q\phi$ or $QG \rightarrow Q\phi$. After the quark-hadron transition, in the absence of free quarks, the dominant production mechanism is nucleon-pion scattering or pion-PNGB conversion, $N\pi \rightarrow N\phi$.

In the case of PNGB coupling only to particles with masses scaled by Λ (case (II) above), we will assume the same processes occurring for the “scaled” matter (and also a transition from scaled quarks to scaled hadrons at $T \sim \Lambda$): $\tilde{Q}\gamma \rightarrow \tilde{Q}\phi$ and $\tilde{N}\tilde{\pi} \rightarrow \tilde{N}\phi$, where all fermion masses have been scaled by $\lambda \simeq \Lambda/\Lambda_{\text{QCD}}$, *e.g.*, $m_{\tilde{Q}} \equiv \lambda m_Q \simeq (\Lambda/\Lambda_{\text{QCD}})m_Q$.

For PNGB photoproduction, the production rate $\Gamma = n_Q \langle \sigma |v| \rangle \sim \alpha T (m_f/f)^2$, and $H \sim T^2/m_{\text{pl}}$ in a radiation-dominated universe, so

$$\frac{\Gamma}{H} \sim \alpha \left(\frac{m_{\text{pl}}}{T} \right) \left(g_Q \frac{m_Q}{f} \right)^2 \quad T \gtrsim \Lambda. \quad (4.4)$$

For pion-PNGB conversion, the nucleon abundance is Boltzmann-suppressed, $n_N \simeq (m_N T)^{3/2} e^{-m_N/T}$, and the cross-section $\langle \sigma |v| \rangle \sim T^2/(f m_\pi)^2$, yielding

$$\frac{\Gamma}{H} \sim (m_N T)^{3/2} g_N^2 \frac{m_{\text{pl}}}{f^2 m_\pi^2} e^{-m_N/T} \quad T \lesssim \Lambda. \quad (4.5)$$

Due to the exponential decay of the nucleon abundance, the value of Γ/H will reach a maximum value very soon after the quark-hadron transition at $T \sim \Lambda$. When the PNGB couples to ordinary hadrons (case (I)), it is at that time in thermal equilibrium $((\Gamma/H)_{T \sim \Lambda} \gtrsim 1)$ for

$$f \lesssim 6 \times 10^8 g_N \text{ GeV} \quad (\text{I}). \quad (4.6)$$

We will define $x = m_N/\Lambda$; for normal matter couplings, $x \simeq 5$. For scaled couplings (case (II)), x will differ; *e.g.*, in an $SU(N)$ technicolor model, the technicolor scale $\Lambda \propto \Lambda_{\text{QCD}}/\sqrt{N}$ and the mass of the lightest technibaryon $m_{\text{LTB}} \propto m_N \sqrt{N}$ so

$x = (x_{\text{QCD}}/N_{\text{QCD}})N \simeq 5N/3$. In terms of x , the rate for scaled matter is maximized when

$$\frac{\Gamma}{H} \sim 5.2 \times 10^{20} g_N^2 \lambda \left(\frac{f}{\text{GeV}} \right)^{-2} x^{-3/2} e^{-x}. \quad (4.7)$$

where $\lambda = \tilde{m}/m = (\Lambda/\Lambda_{\text{QCD}})(N/3)$ for $SU(N)$ technicolor models. In this case, the requirement for thermal equilibrium becomes, defining $y = x/5$,

$$f \lesssim 6 \times 10^8 g_{\tilde{N}} \text{ GeV } \lambda^{1/2} y^{-3/4} e^{-5(y-1)/2} \quad (\text{II}) \quad (4.8)$$

or, for $x = 5$, $N = 3$,

$$f \lesssim 10^9 (\Lambda/\text{GeV})^{1/2} g_{\tilde{N}} \text{ GeV} \quad (\text{II}). \quad (4.9)$$

When these constraints are combined, the result is shown in Fig. 2. $\Omega_\phi h^2$ is greater than 1 for the region above the appropriate lines; the hatched marks denote the intersection of the regions for which the PNGB is in thermal equilibrium and for which $\Omega_\phi h^2 > 1$. For the axion, which couples to normal matter, the constraint given by Eqn.(4.6) applies, and requires that $f \lesssim 6 \times 10^8 \text{ GeV}$ for the axion to be in thermal in equilibrium. In order for thermal axions to make up a considerable fraction of the closure density, the value of f would have to be quite small, $f \lesssim 5 \times 10^4 \text{ GeV}$. Such a value of f would result in a very strongly coupled axion, already ruled out on astrophysical grounds (see section VI).

V. Unstable Particles

Thus far, we have assumed that the PNGB is stable (at least over times longer than the age of the universe). If it decays with a lifetime τ into products which are still relativistic at the present time, the original PNGB density will have been redshifted away between the time of decay and today:

$$\Omega_{\text{Decay Products}}(t_0) = n! \frac{R(\tau)}{R_0} \Omega_{\text{Stable}} \quad (5.1)$$

where R is the scale factor of the universe, R_0 is its present value, n is such that $R \propto t^n$ at the time of decay, $t \sim \tau$, Ω_{Stable} is the density the species would have if it had not decayed, and Ω_{DP} is the present density of the decay particles.

If the particle decays when the universe is matter-dominated ($\tau \gtrsim t_{\text{EQ}} = 4.4 \times 10^{10} (\Omega_0 h^2)^{-2} \text{ sec}$), $n = 2/3$, while if it decays in the radiation-dominated era ($\tau \lesssim t_{\text{EQ}}$), $n = 1/2$. During the matter-dominated epoch, $R(t)/R_0 = 2.9 \times 10^{-12} (\Omega_0 h^2)^{1/3} (t/\text{sec})^{2/3}$. During the radiation-dominated era, $R(t)/R_0 = 2.4 \times 10^{-10} g_*^{-1/12} (t/\text{sec})^{1/2}$. Expressing this in more appropriate units,

$$n! \frac{R(\tau)}{R_0} = \begin{cases} 2.0 \times 10^{-28} \text{ GeV}^{2/3} (\Omega_0 h^2)^{1/3} \tau^{2/3} & \tau > t_{\text{EQ}}; \\ 1.7 \times 10^{-22} \text{ GeV}^{1/2} g_*^{-1/12} \tau^{1/2} & \tau < t_{\text{EQ}}; \end{cases} \quad (5.2)$$

which we parametrize as

$$n! \frac{R(\tau)}{R_0} \equiv A_n \tau^n. \quad (5.3)$$

On dimensional grounds the PNGB decay rate should go as¹ $\Gamma \sim m_\phi^3/f^2$, so the lifetime $\tau = 1/\Gamma$ would be given by

$$\tau = g_d \frac{f^5}{\Lambda^6} = 6.6 \times 10^{17} g_d \text{ sec} \left(\frac{f}{10^{12} \text{ GeV}} \right)^5 \left(\frac{10^3 \text{ GeV}}{\Lambda} \right)^6 \quad (5.4)$$

where g_d is a dimensionless coefficient of order one. (For comparison, the present age of the universe $t_0 = f(\Omega)H_0^{-1} = 4 \times 10^{17} \text{ sec}$ for $\Omega = 1$, $h = 1/2$.) This gives

$$\Omega_{\text{DP}}(t_0) = A_n g_d^n f^{5n} \Lambda^{-6n} \Omega_S(t_0). \quad (5.5)$$

We can now insert our previous expressions for Ω_S , eqns. (2.8), (3.5), and (4.3), and these can then be solved for Λ . For example, PNGBs produced via the misalignment or defect mechanisms have a stable density given by equations (2.8) and (3.5):

$$\Omega_{\phi, \text{mis}} = C_\nu \Lambda f^p; \quad p = \frac{\nu + 3}{\nu + 2}, \quad (5.6)$$

where C_ν is the coefficient of Λf^p in equations (2.8) or (3.5) and depends upon ν , α , g_{*1} , etc. Using this in our results for the density of decay products gives

$$\Omega_{\text{DP}} h^2 = A_n C_\nu g_d^n f^{5n+p} \Lambda^{1-6n} \quad (5.7)$$

or, solving for Λ ,

$$\Lambda = (A_n^{-1} C_\nu^{-1} g_d^{-n} f^{-5n-p} \Omega_{\text{DP}} h^2)^{\frac{1}{1-6n}}. \quad (5.8)$$

In order to convert this expression, and similar ones for the production of PNGBs via string decay and thermal mechanisms, into limits on the PNGB parameters, we need to find a limit on Ω_{DP} . Clearly, requiring $\Omega_{\text{DP}} h^2 < 1$ would be the most conservative bound. However, if we make some assumptions about the nature of these decay products and the growth of structure in the universe, we can put far stricter bounds upon Ω_{DP} .

Observations of the cosmic microwave background¹¹ have determined that its deviations from a perfect blackbody are remarkably small. Specifically, spectral distortions due to the cosmological Sunyaev-Zel'dovich mechanism are below the level of current measurements. When a batch of photons are dumped into the universe (due to a process like PNGB decay), the primary photons will Compton scatter off of free electrons, and the resulting hot electrons will in turn Compton scatter the microwave background, moving photons from the Rayleigh-Jeans to the Wien region. Clearly, this process is only relevant when the photons are produced before the era of recombination: ($\tau \lesssim t_{\text{rec}}$). The amount

of distortion is quantified by the Compton y -parameter $y \cong (1/4)(\Delta\rho_\gamma/\rho_\gamma)$, where $\Delta\rho_\gamma$ is the excess photon density injected by the decays; COBE measurements¹¹ have found that $y \lesssim 3 \times 10^{-4}$. (Actually, at temperatures $T \gtrsim 10^4$ K, the spectral distortion takes the form of a non-zero chemical potential μ , but the resulting bound is similar; we therefore focus on y .) Once the excess photons are dumped into the universe, they evolve the same way as the already-present background photons— y does not vary with time and is proportional to the ratio of Ω_{DP} to Ω_γ .

We parametrize the amount of photon energy dumped into the universe by PNGB decay as $\Delta\rho_\gamma(\tau) = B_\gamma f_c \rho_\phi(\tau)$, with B_γ the branching ratio for decay to photons, and f_c the efficiency of Compton rescattering at the time of PNGB decay ($f_c \simeq 1$ for $T \gtrsim T_{\text{rec}}$, $f_c \simeq 0$ for $T \lesssim T_{\text{rec}}$). Using this for PNGBs with a lifetime $\tau < t_{\text{rec}} \simeq 5.6 \times 10^{12} (\Omega_0 h^2)^{-1/2}$ sec, we find

$$y = \frac{1}{4} B_\gamma f_c \frac{\Omega_{\text{DP}} h^2}{\Omega_\gamma h^2} \lesssim 3 \times 10^{-4} \quad (5.9)$$

or, since $\Omega_\gamma h^2 = 2.6 \times 10^{-5}$,

$$\Omega_{\text{DP}} h^2 \lesssim \frac{3 \times 10^{-8}}{B_\gamma f_c}. \quad (5.10)$$

Although this limit on the CMB distortion via the Sunyaev-Zel'dovich effect is well-understood, it is only effective ($f_c \simeq 1$) for particles decaying at redshifts $10^7 \gtrsim z_{\text{D}} \gtrsim z_{\text{rec}} \simeq 10^3$ *i.e.*, with lifetimes $10^6 \text{ sec} \lesssim \tau \lesssim t_{\text{rec}}$; photons injected at earlier times, $t \lesssim 10^6 \text{ sec}$, will thermalize and produce no distortions. This bound also relies upon assumptions about the PNGB decay branching ratio to photons, B_γ . Because of these constraints, this bound actually gives us no new information, since the area of parameter space with the proper PNGB lifetime does not overlap with the area that is newly allowed by eq. (5.10).

However, we can get a bound on Ω_{DP} independent of the decay products, so long as they are relativistic, by considering the formation of structure²². We know that the universe is matter-dominated today (as long as there are no exotic relativistic particles which came to dominate at very recent epochs; this caveat includes the decay products of the PNGBs, so we cannot put this bound upon Ω_{DP} if $\tau \sim t_0$), and that it must have been matter-dominated for long enough to allow structure formation to occur, since only in a matter-dominated epoch do density perturbations grow sufficiently rapidly, linearly with the scale factor of the universe. The epoch of matter domination is given by

$$1 + z_{\text{eq}} = R_0 / R_{\text{EQ}} = \Omega_{\text{NR}} / \Omega_{\text{rel}} \quad (5.11)$$

where Ω_{rel} is the density of relativistic particles, $\Omega_{\text{rel}} = \Omega_{\gamma+\nu} + \Omega_{\text{DP}}$, and Ω_{NR} is the present energy density of non-relativistic species. Thus, a significant Ω_{DP} will decrease the redshift of matter-radiation equality. Moreover, if $\Omega_{\text{DP}} > \Omega_{\gamma+\nu}$, then the PNGBs dominated the

universe at the time of their decay, and the universe went through two periods of matter-domination: a first phase dominated by the PNGBs before they decayed into relativistic particles, and a second phase dominated by the presently non-relativistic matter. At the final epoch of matter-radiation equality, the comoving wavelength of the horizon scale is (assuming three light neutrino species)

$$k_{\text{eq}}^{-1} = 5 \text{ Mpc} (\Omega_{\text{NR}} h^2)^{-1} \left(1 + \frac{\Omega_{\text{DP}}}{\Omega_{\gamma+\nu}} \right)^{1/2}. \quad (5.12)$$

If the non-relativistic matter today is cold dark matter, k_{eq}^{-1} approximately sets the scale where the present fluctuation spectrum makes the transition from the primordial spectrum on large scales, $|\delta_k|^2 \sim k^n$ for $k^{-1} \gg k_{\text{eq}}^{-1}$, to the processed spectrum on small scales, $|\delta_k|^2 \sim k^{n-4}$ for $k^{-1} \ll k_{\text{eq}}^{-1}$. Enlarging this scale by having an appreciable Ω_{DP} will increase the large scale power in $\delta\rho/\rho$, for a fixed small-scale normalization ($\delta\rho/\rho \approx 1$ at a scale of $8h^{-1} \text{ Mpc}$ today, modulo biasing). The extra large scale power will in turn increase the quadrupole anisotropy of the CMB, which has been constrained by the Relikt experiment²³ to be $(\Delta T/T)_{rms, l=2} < 1.5 \times 10^{-5}$ at the 95% confidence level. (This bound assumes a scale-invariant (Harrison-Zel'dovich) primordial perturbation spectrum; the spectrum-independent bound is roughly a factor of two higher.)

Recently, the qualitatively similar effects of a decaying 17 keV neutrino on structure formation²⁴ have been analyzed. For a primordial scale-invariant spectrum, the quadrupole bound limits its lifetime to be $\tau \lesssim 10 \text{ yr}$; we can translate this constraint into the requirement that²⁵

$$k_{\text{eq}}^{-1} \lesssim 54 h^{-1} \text{ Mpc}. \quad (5.13)$$

Using this in equation (5.12) with $\Omega_{\text{NR}} \approx 1$, we then obtain a limit upon Ω_{DP} :

$$\Omega_{\text{DP}} h^2 \lesssim 115 h^2 \Omega_{\gamma+\nu} h^2 = 5 \times 10^{-3} h^2, \quad (5.14)$$

or, for $h = 1/2$,

$$\Omega_{\text{DP}} h^2 \lesssim 1.3 \times 10^{-3}. \quad (5.15)$$

This corresponds roughly to the requirement that the universe become matter dominated before or during the epoch of recombination. This constraint on Ω_{DP} from structure formation is shown in Fig. 3, for each of the production mechanisms examined above. Decay processes make new areas of the parameter space “cosmologically allowed”: as we move away from $\tau = t_0$ towards shorter and shorter lifetimes (*i.e.*, earlier and earlier decay times), the density of PNGB decay products decreases because it is redshifted away. Note that the region of allowed parameter space for thermally-produced decaying PNGBs, especially those that couple to normal matter, is extremely small; any change in the $O(1)$

parameters that control their abundance is likely to make this region even smaller or eliminate it entirely. Furthermore, we have throughout assumed that each limit applies independently to each production mechanism. If, however, more than one mechanism is effective, then the limits will be tighter, due to separate contributions to Ω .

VI. Astrophysical Constraints

As we have already mentioned, the strength of the coupling of the PNGB to matter and radiation does not depend on the energy scale Λ , but is merely proportional to $1/f$. In fact, any work that restricts the couplings of any pseudoscalar particle is directly applicable to our general PNGBs. Astrophysical methods have been extremely valuable in constraining the mass of one pseudoscalar, the axion, and we will generalize the most important of these results here.

If a PNGB couples strongly enough to matter or radiation, then it will be produced in copious amounts in astrophysical objects such as ordinary stars, red giant stars, and supernovae—if it is also light enough that its production is not Boltzmann-suppressed (m_ϕ less than a few times the temperature of the object). If its interactions are still sufficiently weak, however, it may be able to stream freely out of the star after its production. This “cooling” mechanism—a misnomer, since the necessity of maintaining an equilibrium configuration generally increases the temperature—may strongly affect the evolution of these objects.

The strongest stellar-evolutionary bounds to pseudoscalar couplings come from the examination of helium-burning (horizontal branch) stars. The coupling to photons is constrained by the lifetime of these stars, observed via the fraction of stars seen in that area of a star cluster’s color-magnitude diagram. In the presence of PNGB cooling via the PNGB Primakoff process, $\gamma + (Z, e^-) \rightarrow \phi + (Z, e^-)$, the length of this evolutionary stage will be shortened; a value of

$$g_{\phi\gamma\gamma} \lesssim 10^{-10} \text{ GeV}^{-1} \quad \text{or} \quad f \gtrsim 10^7 g_\gamma \text{ GeV} \quad (6.1)$$

is necessary to keep the helium burning lifetime within an order of magnitude of observations. This constraint holds for PNGB masses $m_\phi \ll T$. If the PNGB mass is greater than the core temperature for typical horizontal branch stars $T \simeq 8.6 \text{ keV}$, the PNGB emission rate will be suppressed by roughly the factor $\exp(-m_\phi/T)$; this ignores any other mass-dependence of the emission rates, which will be swamped by the Boltzmann factor for $m_\phi \gg T$. (We have worked out the limits from Compton emission in horizontal branch stars using the exact temperature-dependence of the rate from Raffelt and Starkman²⁶ and the constraints on f and Λ make only a small difference in fig. 4 below.) In general,

emission rates are proportional to $g_{\phi\gamma\gamma}^2$, so the constraint will become

$$g_{\phi\gamma\gamma} \lesssim 10^{-10} e^{m_\phi/17.2 \text{ keV}} \text{ GeV}^{-1} \quad \text{or} \quad f \gtrsim 10^7 e^{-m_\phi/17.2 \text{ keV}} g_\gamma \text{ GeV} \quad (6.1a)$$

valid for either $m_\phi \ll 8.6 \text{ keV}$ or $m_\phi \gg 8.6 \text{ keV}$.

The constraint on the coupling to electrons relies on the physics of helium ignition in low-mass red giants. These stars have cores of helium nuclei and degenerate electrons. Due to this degeneracy the pressure does not increase as the temperature rises: when it rises sufficiently for helium burning (via the triple-alpha process) a thermal runaway occurs until the temperature is sufficient for degeneracy to be lifted. If PNGBs couple sufficiently strongly to the degenerate electron gas, they will be able to cool the core efficiently enough to prevent the ignition of helium—and this time the term “cooling” is appropriate as the temperature is decreased due to the PNGB streaming. The PNGBs are produced largely via the process of PNGB bremsstrahlung, $e^- + Ze \rightarrow e^- + Ze + \phi$ with an internal photon line. Calculations²⁶ show that this cooling will prevent helium burning for $g_{\phi ee} \gtrsim 3 \times 10^{-13}$. In order for the cooling to actually be effective, however, the PNGBs cannot couple so strongly that they remain trapped in the core, keeping the temperature high. This results in a “window” for which helium ignition will *not* occur:

$$3 \times 10^{-13} \lesssim g_{\phi ee} \lesssim 6 \times 10^{-7}; \quad (6.2)$$

that is, helium ignition will occur normally for

$$f \lesssim 10^3 g_e \text{ GeV} \quad \text{or} \quad f \gtrsim 10^9 g_e \text{ GeV}. \quad (6.3)$$

In order for these bounds to hold, the mass of the PNGB must be much less than the temperature. For heavy PNGBs, the cooling limit is modified as in equation (6.1a). In the trapping regime, the primary effect of masses $m_\phi \gg T$ will be to slow the PNGBs to non-relativistic velocities, decreasing the effective coupling $\alpha_\phi = g_{\phi ee}^2/4\pi$ by a factor of the average thermal velocity of the PNGBs, $\alpha'_\phi = \alpha(m_\phi/T)^{1/2}$, thus reducing the coupling constant $g_{\phi ee}$ by a factor of $(m_\phi/T)^{1/4}$. In order to interpolate between the limits of $m_\phi \ll T$ and $m_\phi \gg T$, we will modify this factor to $(1 + m_\phi/T)^{1/4}$. Thus our Helium ignition limits become

$$f \lesssim 10^3 (1 + m_\phi/8.6 \text{ keV})^{1/4} g_e \text{ GeV} \quad \text{or} \quad f \gtrsim 10^9 e^{-m_\phi/17.2 \text{ keV}} g_e \text{ GeV}. \quad (6.3a)$$

The strictest astrophysical bounds on pseudoscalar couplings arise from observations of the duration of the neutrino pulse from supernova 1987a. If the nucleon coupling is strong enough, PNGB emission will cool the supernova core so efficiently that the neutrino burst duration will be shortened. If they couple too strongly, however, PNGBs will be trapped

in a “PNGB sphere” (analagous to the neutrino-sphere) which will again lengthen the neutrino burst sufficiently to coincide with observation. The dominant emission process in the SN core is PNGB-Nucleon bremsstrahlung, $NN \rightarrow NN + \phi$ with an internal pion line.

Numerical and analytical studies²⁶ have shown that the neutrino pulse will be unacceptably shortened for

$$10^{-10} \lesssim g_{\phi NN} \lesssim 2 \times 10^{-7} \quad (6.4)$$

assuming equal coupling to neutrons and protons ($g_{\phi nn} = g_{\phi pp} = g_{\phi NN}$). Thus the neutrino burst will occur normally for

$$f \lesssim 5 \times 10^6 g_N \text{ GeV} \quad \text{or} \quad f \gtrsim 10^{10} g_N \text{ GeV}. \quad (6.5)$$

If we take account of masses $m_\phi \gg T_{\text{core}} \sim 30 \text{ MeV}$ as above, these constraints become approximately

$$f \lesssim 5 \times 10^6 (1 + m_\phi/30 \text{ MeV})^{1/4} g_N \text{ GeV} \quad \text{or} \quad f \gtrsim 10^{10} e^{-m_\phi/60 \text{ MeV}} g_N \text{ GeV}. \quad (6.5a)$$

Of course, none of these astrophysical bounds obtain unless the PNGB couples to normal matter and electromagnetic radiation. If, for example, the PNGB couples only to the photon, then only the bound on $g_{\phi\gamma\gamma}$ from the helium burning lifetime will matter. Similarly, in the case of the hadronic axion, which does not couple at tree level to the electron (*i.e.*, g_e is very small) only the photon and nucleon bounds are relevant.

We show the astrophysical bounds in Fig. 4; for modest PNGB masses, the limits are merely upon the parameter f that controls the PNGB couplings. For extremely heavy PNGBs, however, the masses are comparable with the temperatures of the astrophysical objects under consideration, and the limits are strongly curtailed.

For small values of f , the coupling of the PNGB to matter becomes strong enough that it would show up in terrestrial accelerators. On the basis of nondetection of $\Upsilon \rightarrow \phi + \gamma$ and $J/\psi \rightarrow \phi + \gamma$ events, we can limit¹

$$f \gtrsim 10^3 \text{ GeV}.$$

This bound holds for PNGB masses $m_\phi \lesssim 1 \text{ MeV}$, above which the decay to electron-positron pairs becomes important. Thus, the lower left of fig. 4 is also disallowed, despite the regions allowed by trapping in helium-burning stars.

Other groups have also been investigating various astrophysical constraints on low-mass PNGB properties. Gnedin and Mészáros²⁷ have examined the conversion of PNGBs to radio photons in the magnetic field of a pulsar. Engel *et al.*²⁸ have used the nondetection of a direct axion signal from SN87a by the Kamiokande II detector to exclude the range

$9.4 \times 10^2 \text{ GeV} \lesssim f/g_N \lesssim 1.0 \times 10^6 \text{ GeV}$, for $m_\phi \ll 30 \text{ MeV}$. Ressel²⁹ uses the nondetection of a spectral axion line in the night sky resulting from present day decay of clustered axions to exclude $f/g_\gamma \lesssim 2 \times 10^6 \text{ GeV}$, for PNGBs with decay times $\tau \simeq t_0$.

VII. Conclusions

We have considered the cosmological and astrophysical bounds on the parameter space of broken global symmetry models with pseudo-Nambu-Goldstone bosons. As Figs. 1 and 2 indicate, as the soft breaking scale Λ is increased, the cosmological upper bound on the spontaneous symmetry breaking scale f (due to the requirement that PNGBs have less than closure density today) is generally reduced, in accordance with expectations. However, this general trend is subverted when Λ becomes large enough that the PNGB decays on a timescale significantly shorter than the age of the Universe. In that case, since the relativistic PNGB decay products redshift away, the bounds are somewhat relaxed, as Fig. 3 shows. Since the stellar emission of PNGBs depends only on the scale f , the astrophysical lower limits on the axion scale f_{PQ} (from red giants and SN1987a) can be for the most part taken over to the general case. However, these bounds do have an indirect dependence on the scale Λ through the Boltzmann suppression factor for PNGB emission $\sim e^{-m_\phi/T}$, since $m_\phi = \Lambda^2/f$. Thus, for large values of Λ , the astrophysical lower bounds on f are also relaxed. From Fig. 5, which summarizes the astrophysical and cosmological constraints on the parameters Λ and f , we see that there are essentially two allowed regions of parameter space: (i) for $\Lambda \lesssim 1 - 50 \text{ GeV}$ (depending on the DS vs. HS debate) and $f \gtrsim 10^{10} \text{ GeV}$, there is a triangular region allowed for stable, light ($m_\phi \lesssim 10^{-5} \text{ eV}$) PNGBs; here the upper bound on f due to the cosmic PNGB density increases as Λ is reduced. (ii) For $\Lambda \gtrsim 10 - 1000 \text{ GeV}$, there is a second allowed region; as Λ is increased above this limit, a widening window for f opens up. In this region, the PNGBs are heavy ($m_\phi \gtrsim 1 \text{ MeV}$) and unstable on a timescale short compared to the age of the Universe.

Acknowledgements

This research was supported in part by the DOE (at Chicago and Fermilab) and by NASA Grant NAGW-1340 at Fermilab. Parts of this work were completed at the 1991 Aspen Center for Physics summer workshop.

REFERENCES

- [1] For a review, see, *e.g.*, J. E. Kim, *Phys. Rep.* **150**, 1 (1987).
- [2] J. Preskill, M. Wise, and F. Wilczek, *Phys. Lett.* **120B**, 127 (1983); L. Abbott and P. Sikivie, *ibid.* **120B**, 133 (1983); M. Dine and W. Fischler, *ibid.* **120B**, 137 (1983).
- [3] For reviews, see G. Raffelt, *Phys. Rep.* **198**, 1 (1990); M. S. Turner, *Phys. Rep.* **197**, 67 (1990).
- [4] C. T. Hill and G. G. Ross, *Phys. Lett.* **203B**, 125 (1988); *Nucl. Phys.* **B311**, 253 (1988).
- [5] E. Witten, *Phys. Lett.* **149B**, 351 (1984); K. Choi and J. E. Kim, *ibid.* **154B**, 393 (1985).
- [6] J. P. Derendinger, L. E. Ibanez, H. P. Nilles, *Phys. Lett.* **155B**, 65 (1985); M. Dine, R. Rohm, N. Seiberg, and E. Witten, *ibid.* **156B**, 55 (1985).
- [7] K. Freese, J. Frieman, and A. Olinto *Phys. Rev. Lett.* **65**, 3233 (1990).
- [8] M. S. Turner, *Phys. Rev. D* **33**, 889 (1986).
- [9] See, *e.g.*, D. J. Gross, R. D. Pisarski, and L. G. Yaffe, *Rev. Mod. Phys.* **53**, 43 (1981).
- [10] A. Gupta, C. Hill, R. Holman, and E. Kolb, Fermilab preprint (1991).
- [11] J. Mather, *et al.*, *Astrophys. J. Lett.* **354**, 37 (1990). H. P. Gush, M. Halpern, E. H. Wishnow, *Phys. Rev. Lett.* **65**, 537 (1990).
- [12] For discussion and references, see *e.g.*, E. Kolb and M. S. Turner, *The Early Universe*, (Addison-Wesley, New York) (1990).
- [13] M. S. Turner and F. Wilczek, *Phys. Rev. Lett.* **66**, 5 (1991); A. Linde, Stanford preprint SU-ITP-883 (1991).
- [14] T. W. B. Kibble, *Phys. Rep.* **67**, 183 (1980).
- [15] R. Davis, *Phys. Lett.* **180B**, 225 (1986); R. Davis and P. Shellard, *Nucl. Phys.* **B324**, 167 (1989); A. Dabholkar and J. Quashnock, *Nucl. Phys.* **B333**, 815 (1990).
- [16] D. Harari and P. Sikivie, *Phys. Lett.* **195B**, 361 (1987).
- [17] D. Bennett and S. Rhie, *Phys. Rev. Lett.* **65**, 1709 (1990).
- [18] N. Turok, *Phys. Rev. Lett.* **63**, 2635 (1989); Princeton preprint PUPT-90-1230 (1990).
- [19] For a review, see A. Vilenkin, *Phys. Rep.* **121**, 263 (1985); for numerical simulations of string-wall networks, see W. Press, B. Ryden, and D. Spergel, *Astrophys. J.* **347**, 590 (1989).
- [20] P. Sikivie, *Phys. Rev. Lett.* **61**, 783 (1988).
- [21] M. S. Turner, *Phys. Rev. Lett.* **59**, 2489 (1987).
- [22] G. Steigman and M. S. Turner, *Nucl. Phys.* **B253**, 375 (1985).
- [23] A. A. Klypin, M. V. Sazhin, I. A. Strukov and D. P. Skalachev, *Sov. Astr. Lett.* **13**,

104 (1987).

- [24] J. R. Bond and G. Efstathiou, preprint (1991).
- [25] In applying this constraint, we have assumed that the comoving Hubble scale at the time when the PNGBs come to dominate the mass density (*i.e.*, at the onset of the first matter-dominated phase), $k_{\text{eq},1}^{-1}$, is well below the scale at which the spectrum is currently normalized, $k_{\text{norm}}^{-1} \simeq 10h^{-1}$ Mpc. Now, $k_{\text{eq},1}^{-1} = 5 \text{ Mpc } (\Omega_{\phi,s} h^2)^{-1}$, where $\Omega_{\phi,s}$ is the present density the PNGBs would have if they were stable. If $\Omega_{\text{NR}} = 1$, there would be a PNGB-dominated phase only if $\Omega_{\phi,s} > 1$, *i.e.*, only if $k_{\text{eq},1}^{-1} < k_{\text{norm}}^{-1}$. We thus expect corrections to the bound (5.13) due to variations in $k_{\text{eq},1}$ to be small.
- [26] For a review of astrophysical bounds, see Raffelt *op. cit.*³, and references therein; G. Raffelt and G. Starkman, *Phys. Rev. D* **40**, 942 (1989); A. Burrows, M. T. Ressell, M. S. Turner, *Phys. Rev. D* **42**, 3297 (1990); A. Burrows, M. S. Turner, R. Brinkman, *Phys. Rev. D* **39**, 1020 (1989).
- [27] Yu. N. Gnedin and P. Mészáros, preprint (1991).
- [28] J. Engel, D. Seckel, A. C. Hayes, *Phys. Rev. Lett.* **65**, 960 (1990).
- [29] M. T. Ressell, to be published in *Phys. Rev. D* (1991); M. A. Bershadsky, M. T. Ressell, M. S. Turner, *Phys. Rev. Lett.* **66**, 1389 (1991).

FIGURES

FIG. 1. The excluded area of the Λ - f plane for non-thermal misalignment production of stable PNGBs. The region above these lines has $\Omega_\phi h^2 > 1$ and is excluded. The line labeled “base” is the axionic case, with the following parameters: $\nu = 3.7$, $\alpha = 0.1$, $g_{*1} = 100$, $\theta_1 = \pi/\sqrt{3}$. The other lines differ from this case as marked in the legend. In addition, lines of constant mass ($m_\phi = \Lambda^2/f$) are shown for reference.

FIG. 2. The excluded area of the Λ - f plane for all production mechanisms of stable PNGBs. For Misalignment and String production, the area above the $\Omega_\phi h^2 = 1$ lines is excluded as in Fig. 1. For Thermal production, the excluded area is above the line marked “In TE” and to the left of either the line marked “Thermal (Normal)” or “Thermal (Scaled),” depending on the PNGB couplings (see text). In this figure and below, we have set $g_N = 1$.

FIG. 3. The excluded area of the Λ - f plane for decaying PNGBs, limited by constraints of structure formation and CMB distortion, $\Omega_{\text{DP}} h^2 \lesssim 1.3 \times 10^{-3}$. The lines are as in Fig. 2, with the constraints bending over as the density redshifts away for PNGBs that have decayed by the present time. In this figure, we have set $g_d = 1$.

FIG. 4. The excluded area of the Λ - f plane due to astrophysical limits on PNGB couplings. The regions below the curves are disallowed by the various astrophysical arguments as marked. In this figure, we have set $g_e = g_N = g_\gamma = 1$ (see text). Lines of constant PNGB mass are shown for reference.

FIG. 5. The allowed areas of the Λ - f plane with all constraints from Figs. 3–4 combined. Solid lines are constraints from misalignment and topological defects; dotted lines are from thermal production; short dashed lines are from astrophysical constraints. “PNGB dark matter” labels the region where $\Omega_\phi h^2 \sim 1$. Lines of constant PNGB mass are shown for reference (long dashes).

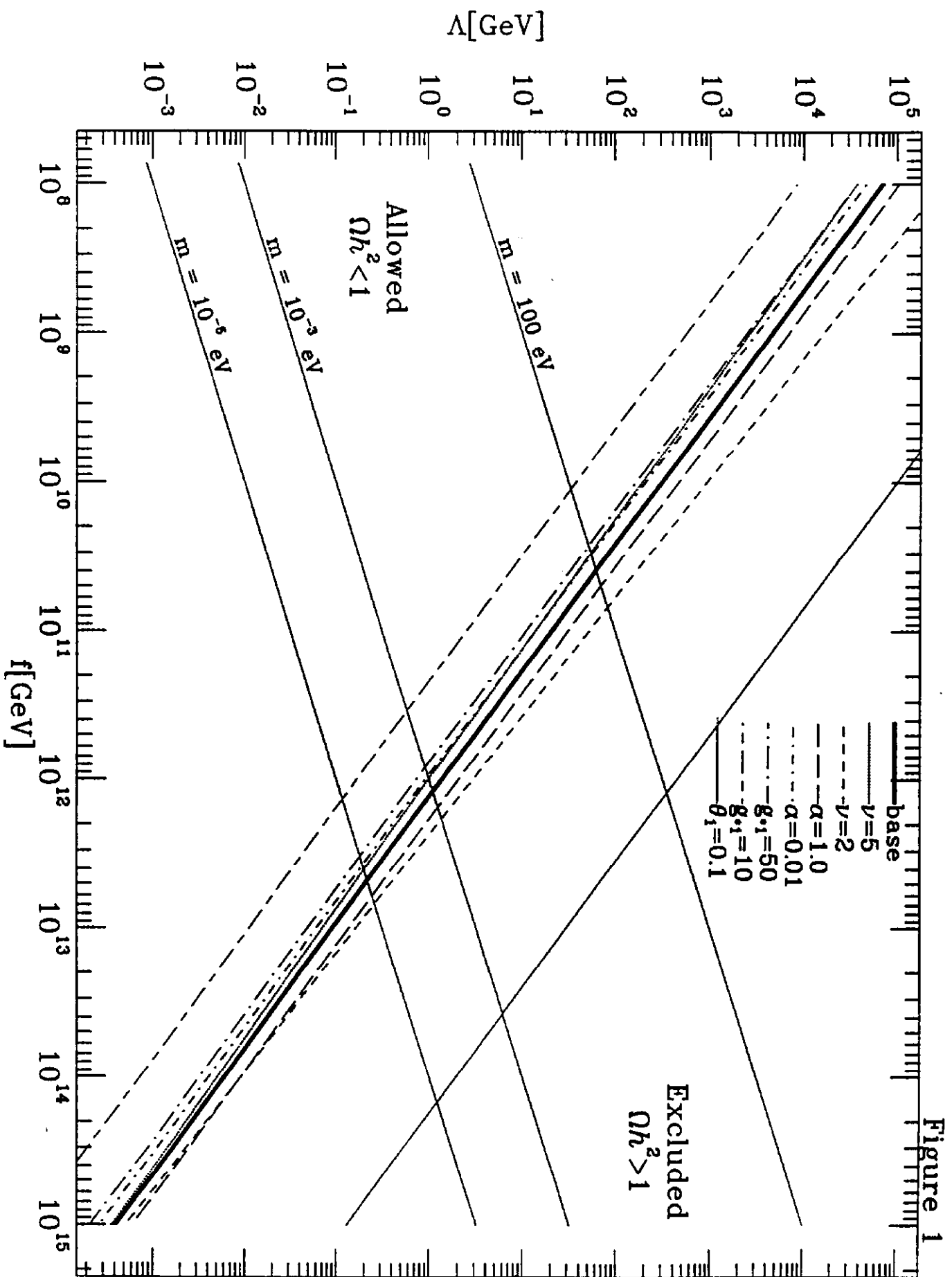


Figure 1

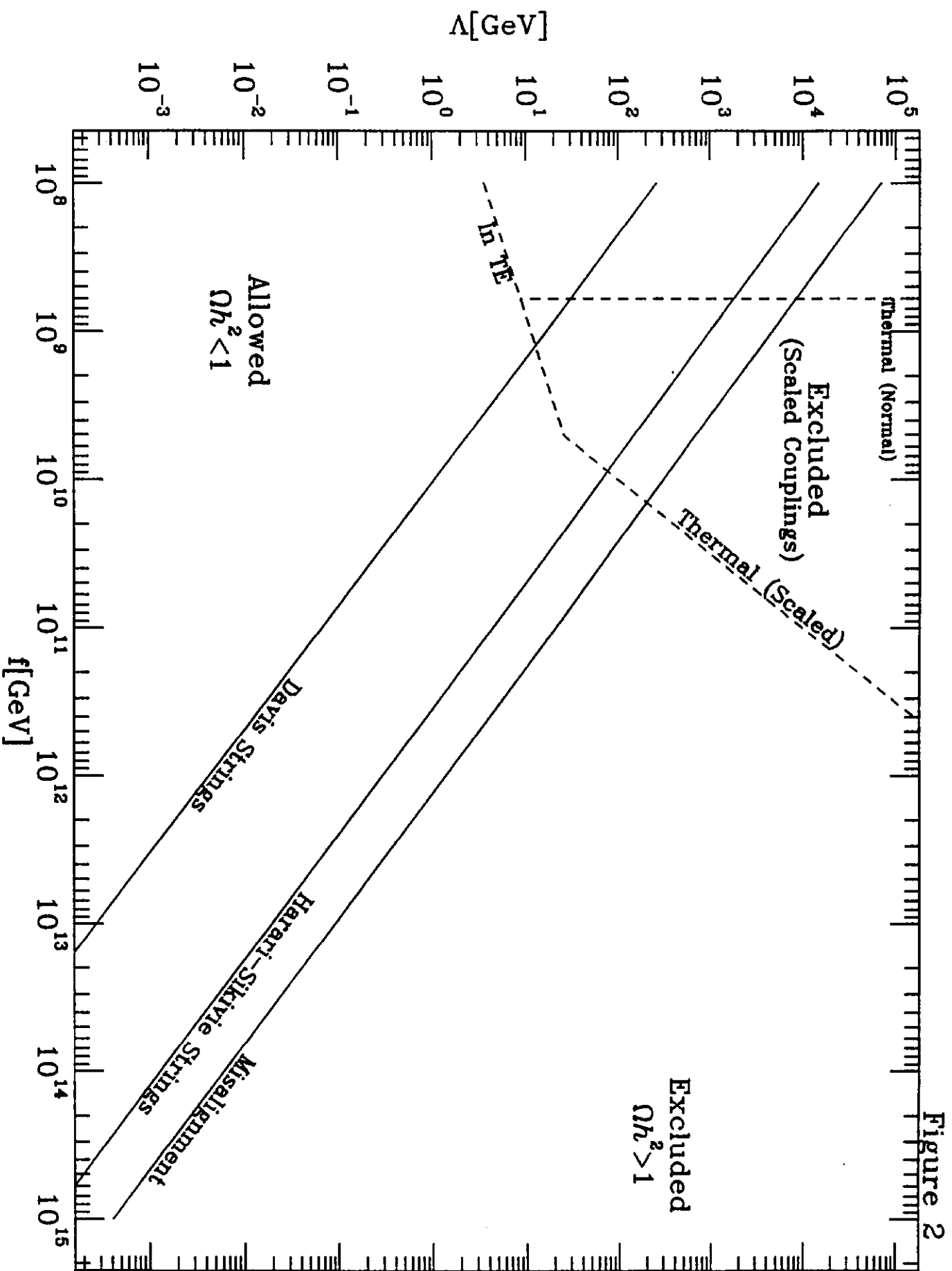


Figure 2

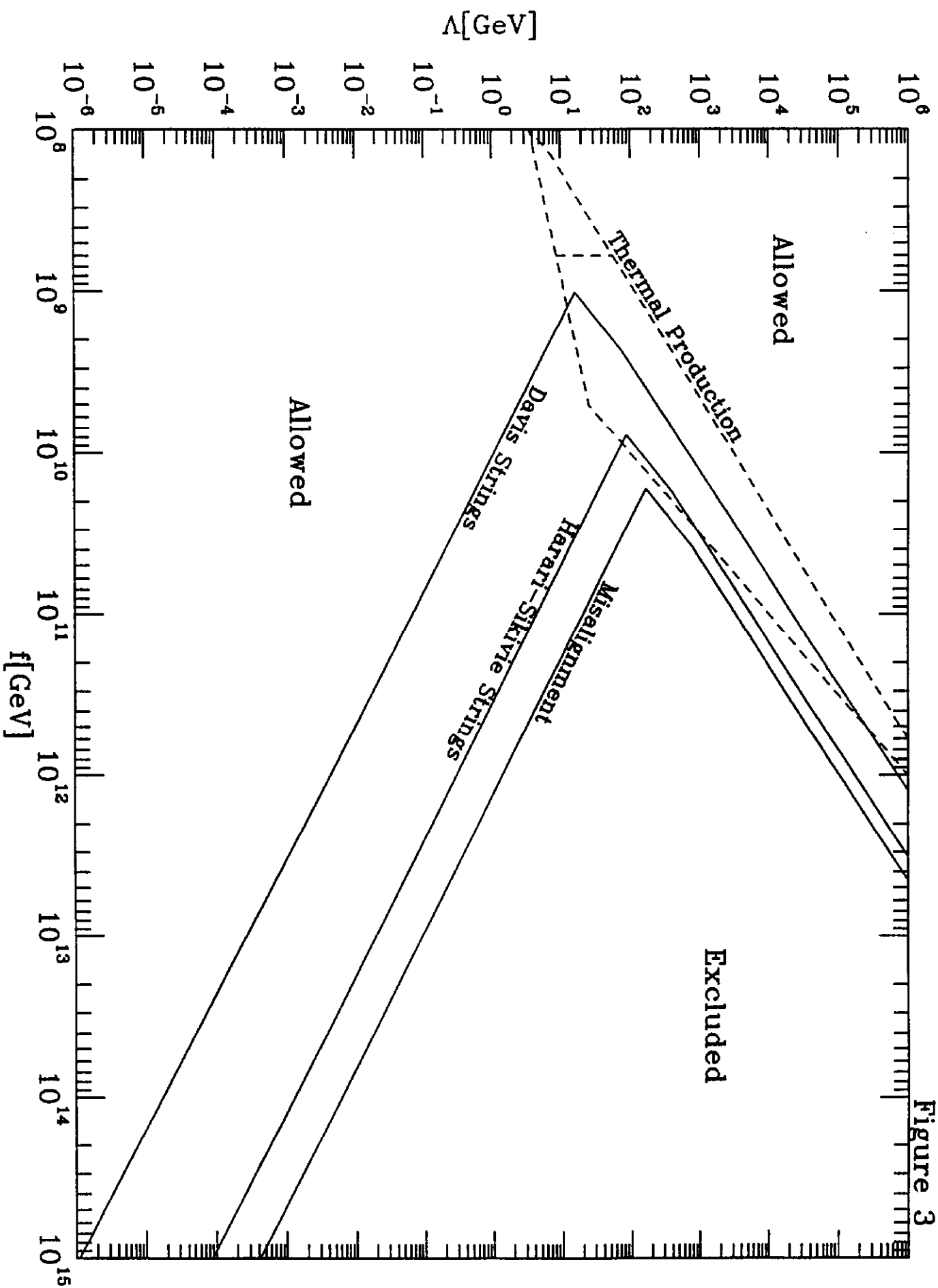


Figure 3

Figure 3

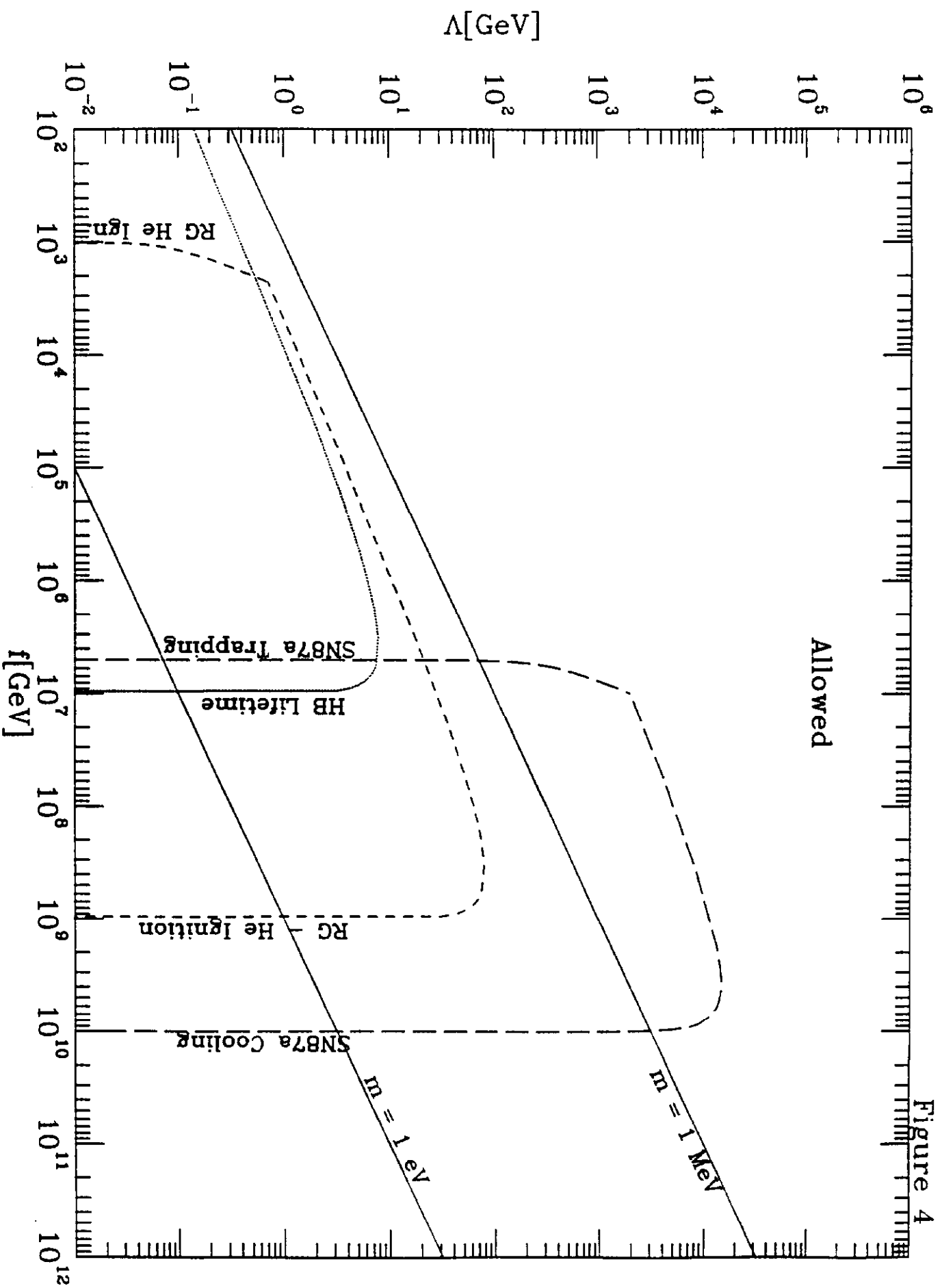


Figure 4

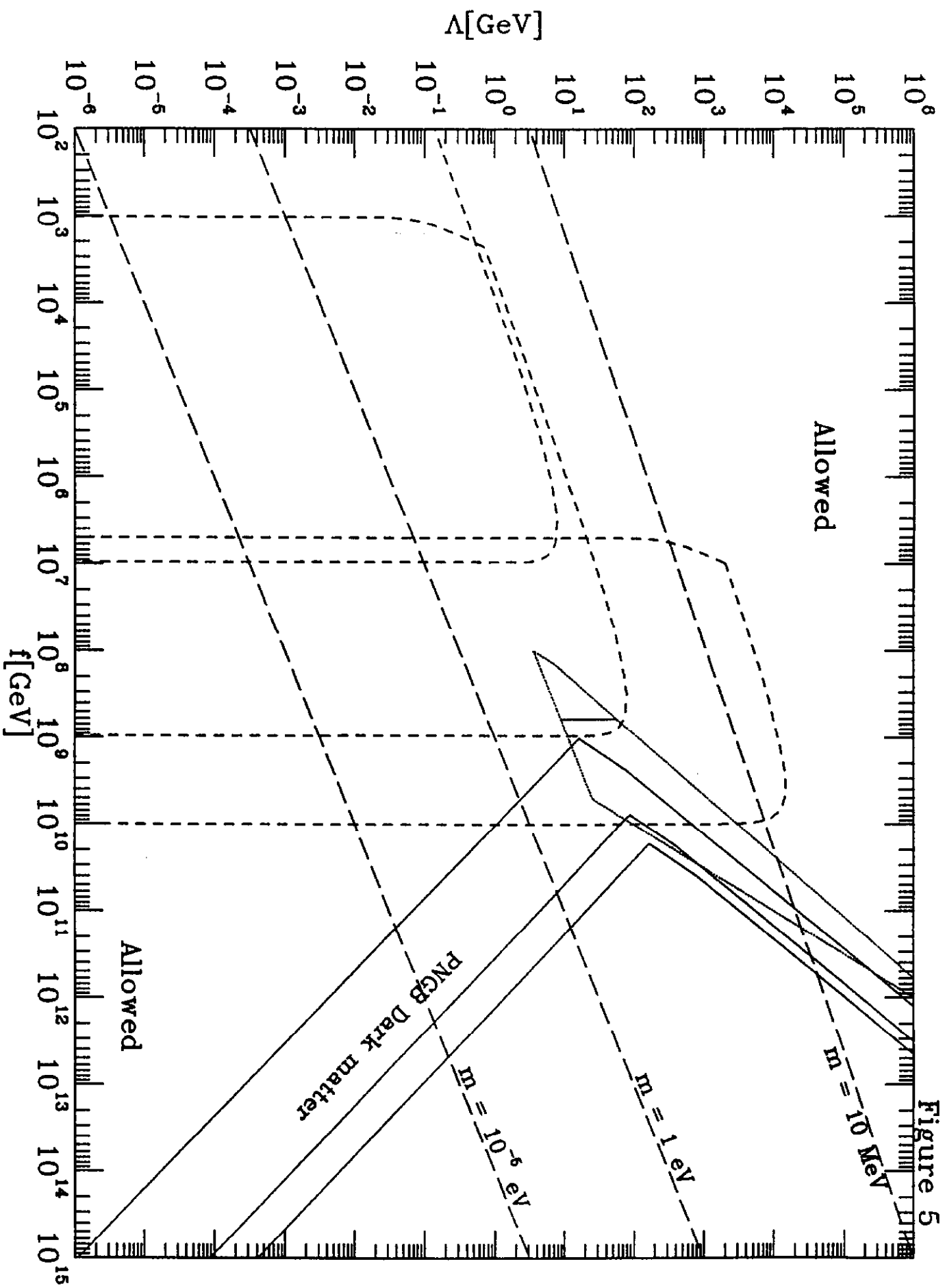


Figure 5

Figure 5